

p. 262 #35, 36, 42, 47 thru 50

42) faculty: randInt(1, 10) --> choose 2

students: randInt(1, 30) --> choose 4

total sample = 6 people

Parameter

- * Describes a POPULATION
- * Is a fixed (set) #
- * often unknown

Statistic

- * Describes a SAMPLE
- * Changes from sample to sample
- * We use repeated values to help estimate a population parameter

Sampling Variability

- * the value of a statistic VARIES in repeated samplings

* Greek letters

μ = avg. ht of HS students

P = % of heads in 10 coin flips

* English letters

\bar{X} = 64.2"

\hat{P} = 3/10 = 30%

$\hat{P}_1 = 30\%$
 $\hat{P}_3 = 40\%$

$\hat{P}_2 = 70\%$
 $\hat{P}_4 = 40\%$

Simulation

- * using random digits to represent events (based on the probability of the events)
- * generating those digits to represent a sample

Example:

Dice: outcomes: #1-6
equal

$\text{randInt}(1, 6)$

Coin toss: H T
0 1
#1-5 6-10

$\text{randInt}(0, 1)$
 $\text{randInt}(1, 10)$

Baseball player:

35% chance of getting a hit

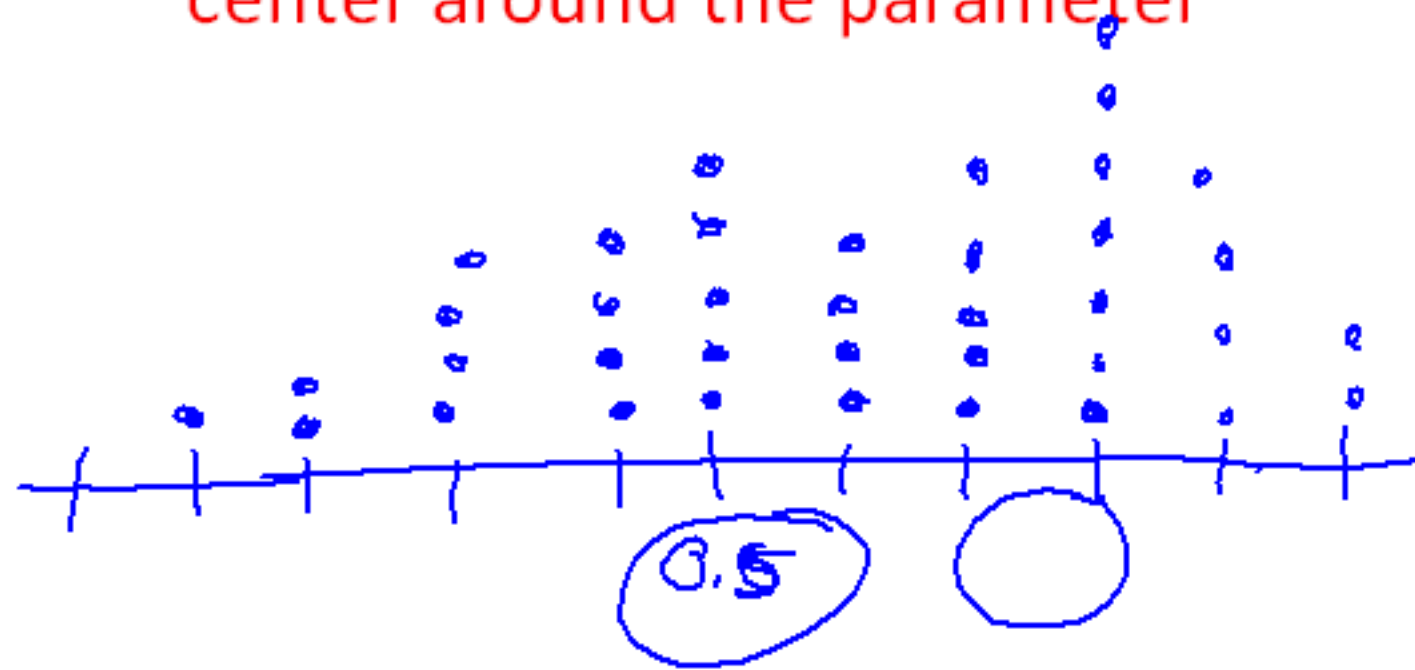
Hit #1-35

Miss #36-100

$\text{randInt}(1, 100)$

Sampling Distribution =

- * values taken by a statistic in repeated samples
- * shown in a histogram
- * should be approximately normal, center around the parameter



\hat{p} = % of heads
in 10 coin
flips

$p = ?$

Unbiased Estimator =

- * When the center of the sampling distribution = the parameter

Variability of a statistic=

* the spread of a sampling distribution

* as $n \uparrow$, spread of distrib. \downarrow

Repeated samples

$n = 30$

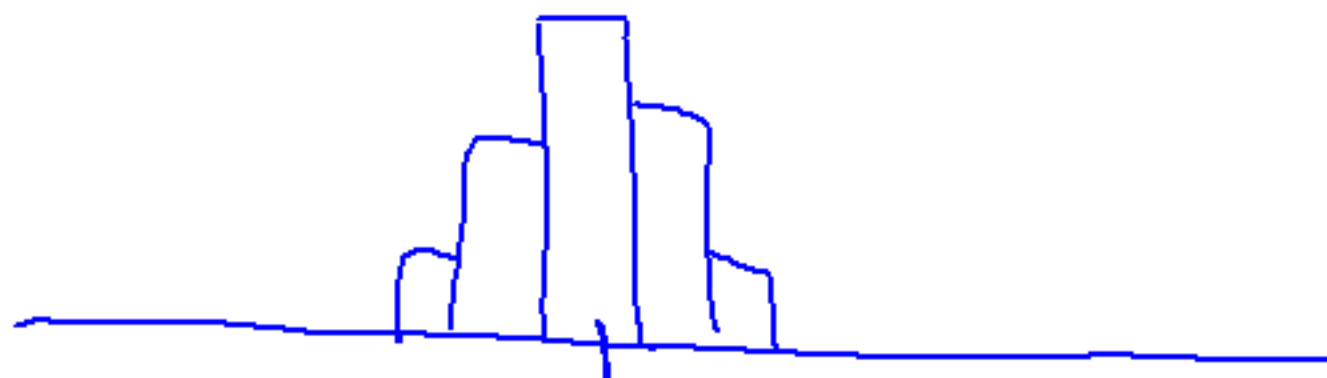


$n = 100$



*

$n = 1000$

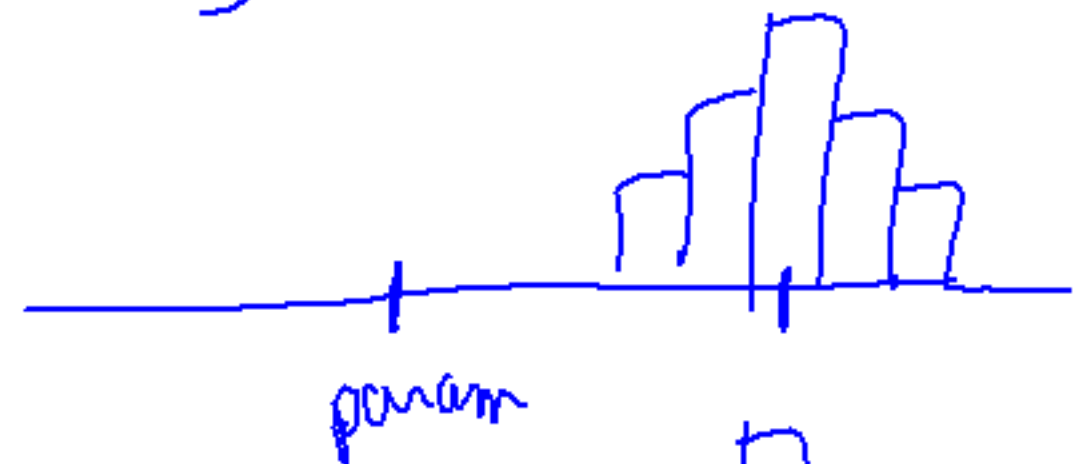


Bias vs. Variability:

pictures p. 274

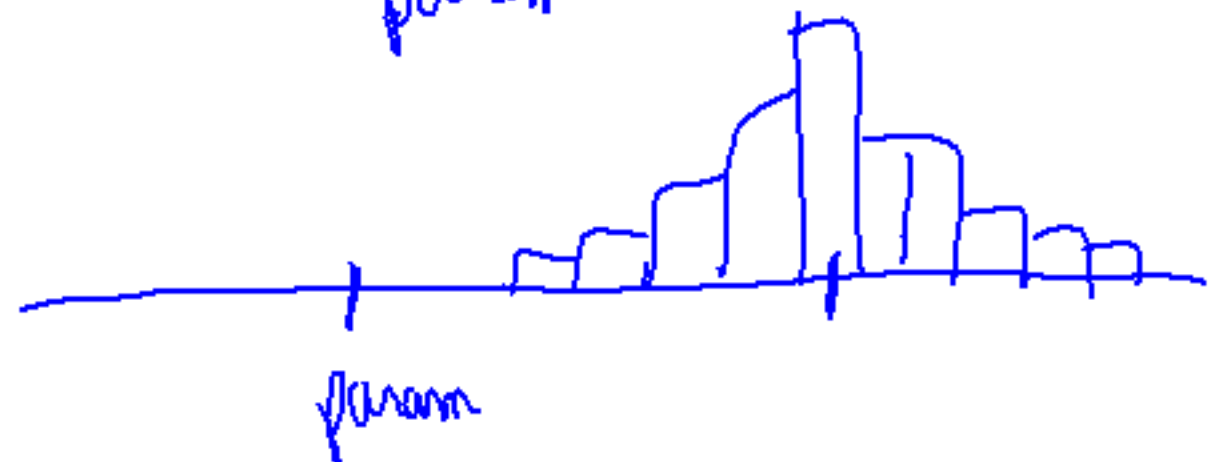
spread/precision

accuracy of
your statistics
(to the true parameter)



Why randomize?

* READ p. 275!



51) 6.2% = statistic

52) 2.503 = parameter

2.515 = statistic

53) 43 = statistic

52% = parameter

54) 335 and 289 = statistics

55) (a) higher bias, high variability

(b) low bias, low variability

(c) low bias, high variability

(d) high bias, low variability

57) (a) NO. since all sample sizes are 2000, the variability will be the same for all states

(b) YES. Since the sample sizes are all different, the variability will all be different.

AP 2008 FORM B #2

Part (a)

A, C and D appear to be unbiased, because the centers of their sampling distributions is the same as the population parameter (75)

Part (b)

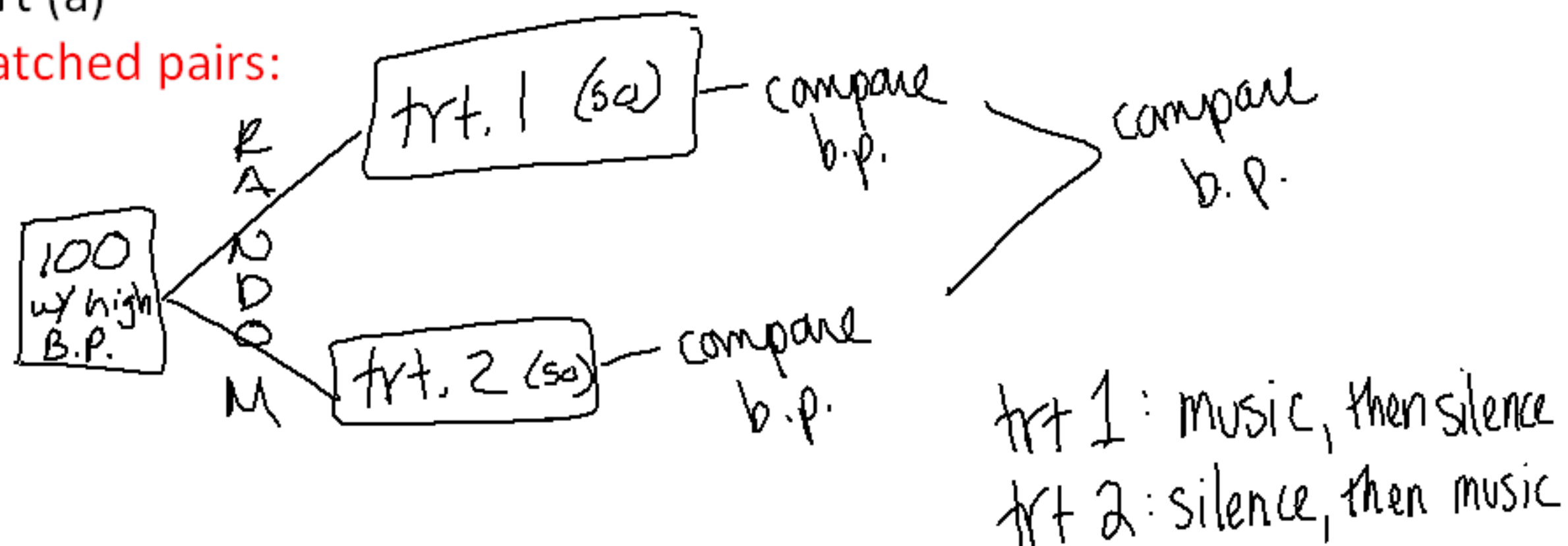
Statistic A would be better because it is unbiased, so it would do a better job estimating the parameter. B is biased, so it would NOT do a good job estimating the parameter.

Part (c)

Statistic C would do a better job because it has lower variability. Since both C and D are unbiased, we must look at the variability of the two. C would give us estimates that are closer to the parameter.

Part (a)

Matched pairs:



CRD:

